

Influencia del nivel educativo y otros co-determinantes en la probabilidad individual de sufrir desempleo.

Navarro Gómez, María Lucía

Rueda Narváez, Mario F.*

Dpto. de Estadística y Econometría. Universidad de Málaga

*autor para correspondencia: mfrueda@gmail.com

El desempleo es uno de los principales problemas de la economía Española, con una evolución del mercado de trabajo notablemente peor en los últimos años que en el resto de países comparables. La tasa de paro española destaca con un 17%, muy por encima de la media europea del 7,6%, y sólo por debajo de Grecia (Eurostat, datos para agosto de 2017). Pese a una evolución reciente positiva tras el máximo del 26% en 2013, la economía sigue lejos de la situación previa de 2007, antes del comienzo de la crisis, en la que el paro bajó hasta el 8,2% y se situaba apenas un punto por encima de la media Europea. Este anormal nivel de desempleo no se traduce en una diferencia tan acusada en cuanto a la tasa de empleo, que en España era a mitad de 2017 de un 61,1%, frente a una media del 66,7% en la UE.

En este contexto, es interesante estudiar hasta qué punto las anormalmente altas tasas de paro en España responden a dificultades generalizadas para encontrar empleo para todos los trabajadores por igual, o si bien existen diferencias entre distintos grupos, de acuerdo a características como el nivel educativo, el género o la edad, entre otras. El presente trabajo analiza cómo las características individuales de los trabajadores afectan a la probabilidad de que estén parados frente a ocupados, y cómo tales efectos han evolucionado desde el comienzo de la crisis en 2008. Para ello, utilizamos datos de la EPA desde 2005 hasta 2012 para estimar modelos logit en los que la variable dependiente es un indicador de desempleo (frente a ocupación), estudiando la evolución de los resultados obtenidos en los distintos trimestres durante los últimos años. Los resultados permitirán identificar los grupos con mayores problemas de desempleo y qué colectivos han encontrado mayores dificultades relativas para acceder al empleo en los últimos años.

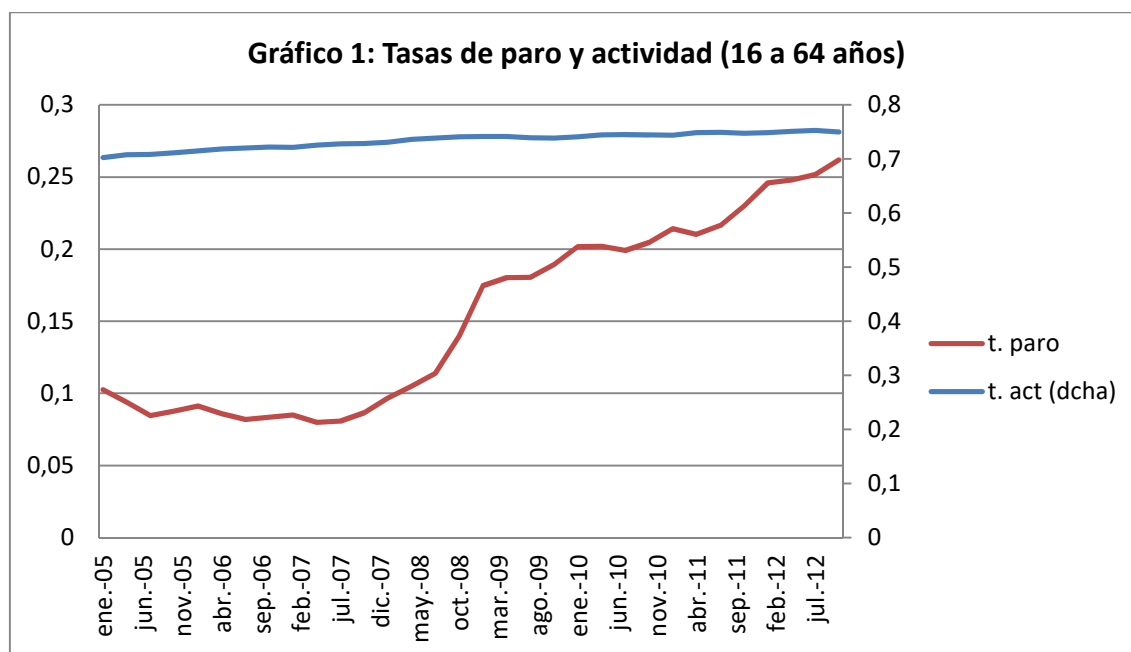
Palabras clave: Desempleo, educación, modelos de variable dependiente limitada

Códigos JEL: J64, J24,

1 Introducción

Una de las consecuencias más preocupantes de la actual crisis económica ha sido la producida en el empleo. La tasa de paro general en España alcanzó un mínimo del 7,95% en el segundo trimestre de 2007 y a partir de ahí creció hasta algo más del 26% a finales de 2012, según los resultados de la Encuesta de Población Activa. Este crecimiento fue en principio extremadamente rápido (hasta principios de 2009) y después más pausado (ver Gráfico 1). La tasa de actividad, mientras tanto, se ha mantenido relativamente estable para una población creciente. En cantidades, hemos pasado de 20,4 millones de ocupados en el segundo trimestre de 2007 a 17 millones a finales de 2012, una pérdida de 3,6 millones de empleos. El número de parados ha aumentado de 1,8 a 6 millones (un incremento de 4,2 millones).

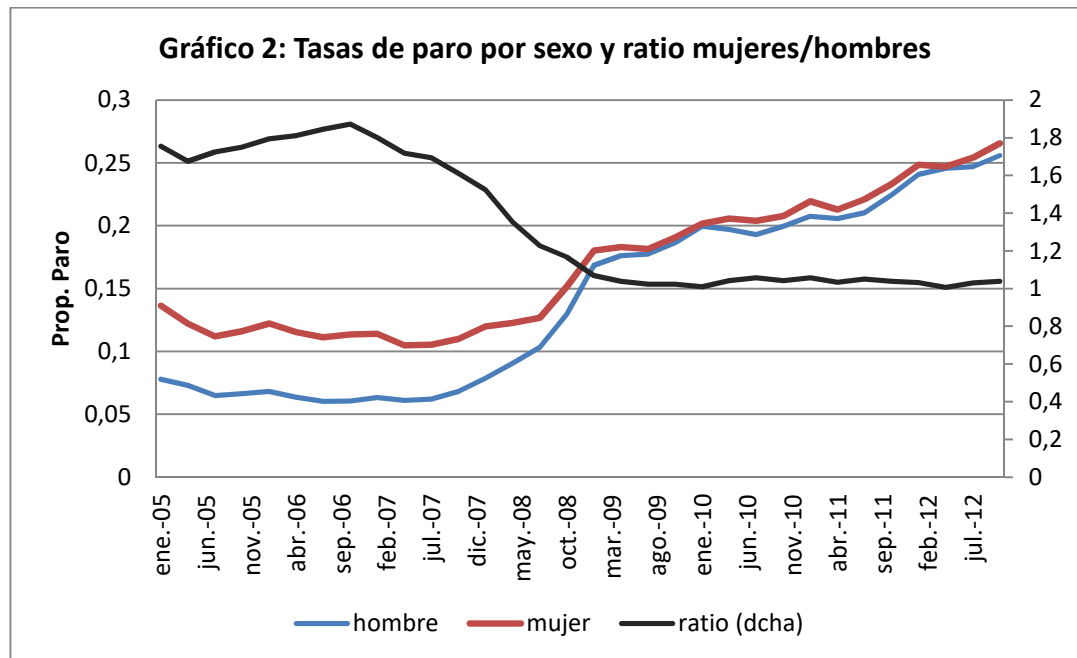
García-Serrano (2011), con datos hasta principios de 2010, presenta un relato resumido del auge y caída de algunos sectores concretos en el empleo, especialmente en la construcción y actividades relacionadas con ella.



Este incremento del desempleo presenta un grave desafío para la economía española y tiene consecuencias graves para el bienestar social, las finanzas públicas, el sistema de pensiones y, posiblemente, los flujos migratorios que experimente el país en el futuro cercano. Este incremento del desempleo presenta un grave desafío para la economía española y tiene consecuencias graves para el bienestar social, las finanzas públicas, el sistema de pensiones y, posiblemente, los flujos migratorios que experimente el país en el futuro cercano. La gravedad específica de la situación en España, con un mayor aumento del paro que en otras economías desarrolladas en el presente ciclo recesivo se explica por la fuerte caída del empleo en sectores muy sobre-representados en España, como la construcción (García-Serrano, 2011). Además, se ha argumentado que la normativa laboral española ha contribuido a crear un mercado de trabajo dual, en el que la alta proporción de trabajadores temporales acumulados en la década pasada ha sufrido, vía despidos o no renovaciones, la necesidad de las empresas de recortar costes ante la caída de la actividad económica (Bentolila y otros, 2012).

Este trabajo tiene como objetivo identificar los principales determinantes individuales del desempleo y analizar su evolución reciente. Para ello, usamos estimaciones *logit* de la probabilidad de estar parado en función de una serie de características personales, del hogar y de la región o provincia de residencia. Los modelos se estiman a partir de los datos de la EPA (2005-2012) y servirán para identificar si el incremento del desempleo ha afectado a todos los grupos sociales por igual o si, por el contrario, existen diferencias en la evolución de la probabilidad de no poder trabajar en función del sexo, el nivel educativo u otras características. Por ejemplo, resulta llamativo que en los años de bonanza la tasa de paro femenina prácticamente duplicara a la masculina, mientras que a partir de 2009 ambas toman prácticamente el mismo valor y comparten la misma tendencia creciente (ver gráfico 2). Ejemplos parecidos de evoluciones dispares pueden encontrarse en función del nivel de estudios (para los trabajadores con educación primaria la probabilidad de estar parado pasa del 10% a casi el 40% entre 2007 y 2012, mientras que aquéllos con estudios universitarios pasan del 5% al 16%) o la edad. Los modelos estimados indicarán cómo afecta cada característica a las posibilidades de encontrarse en paro manteniendo

el resto de las mismas constantes, permitiendo así comprobar si, en términos relativos, algunos trabajadores encuentran dificultades adicionales.



El resto del trabajo se distribuye de la siguiente manera: el próximo epígrafe presenta los datos y el modo de obtención de las variables utilizadas en el análisis, así como la especificación econométrica y el método de estimación. Posteriormente se discuten los resultados en detalle, terminando con las principales conclusiones.

2 Datos y especificación econométrica

Los datos utilizados para este estudio se construyen agregando los ficheros transversales de microdatos individuales de la Encuesta de Población Activa (INE, 2005-2012), desde el primer trimestre de 2005 hasta el último publicado, el correspondiente al cuarto trimestre de 2012. El punto de partida de la serie implica que la base cubre la fase final del periodo expansivo de finales de los '90 y principios de los 2000 hasta la crisis de empleo que comienza a mediados de 2007 y continúa en la actualidad (2013). La amplitud considerada, por tanto, permite estudiar el funcionamiento del mercado de trabajo en ambas fases, expansiva y recesiva.

Originalmente, cada uno de los ficheros trimestrales contiene información sobre unos 160.000 individuos, incluyendo menores. Para nuestro análisis, nos centramos en los activos de menos de 65 años (edad legal ordinaria de jubilación), restringiendo el rango de edades analizadas de 16 a 64 años. Esto nos deja con unos 70.000-80.000 individuos de ambos sexos, cuya información utilizamos para estimar la probabilidad de estar parado (frente a ocupado, al considerar sólo activos) con modelos *logit* para cada ciclo de la encuesta. En el gráfico 1 se muestra la evolución de las tasas de actividad y paro para el rango de edad utilizado, observándose claramente los dos periodos de destrucción intensa de empleo, y no estando claro que la tasa de actividad se haya estabilizado en un máximo (la serie sólo descende un cuarto de punto en el último trimestre de 2012).

Frente a esta situación global, cabe preguntarse si la evolución del desempleo ha sido comparable en distintos grupos de la población activa española de acuerdo a su sexo, nivel educativo y otras características relevantes, lo que constituye el objetivo del presente trabajo. Además del hecho de considerar sólo personas activas (según la metodología de la EPA) y de hasta 64 años de edad, el análisis requiere crear una serie de variables personales sobre las que condicionar la probabilidad de encontrarse en situación de desempleo. Podemos clasificarlas en características de capital humano, personales, familiares y regionales.

En cuanto al capital humano, el nivel educativo más alto alcanzado por la persona se agrupa en 10 categorías, desde aquéllos sin estudios primarios (0) hasta doctorado o máster universitario oficial (10). Además, a fin de aproximar la habilidad de cada individuo con respecto a los estudios formales, creamos un indicador a partir del tiempo que tardó cada individuo en finalizar sus estudios (variable "repetidor"), agrupándolos en 4 categorías¹: Normal (edad de finalización igual o inferior a la moda de la variable), tipo 1 (uno o dos años por encima de la edad más frecuente de finalización), tipo 2 (un intervalo de dos o tres años más, dependiendo de lo dispersa que sea

¹ Este indicador no se crea para los niveles de estudios inferiores a primaria, garantía social, doctorado y cursos superiores de especialización, debido a la naturaleza más heterogénea de estas categorías de estudios.

la edad de finalización) y tipo 3 (por encima de las anteriores). A modo de ilustración, para el nivel de estudios de Bachillerato, consideramos "normal" una edad de finalización hasta 18 años, repetidor tipo 1 hasta 20, tipo 2 hasta 23 y tipo 3 por encima de 23 (un 6,4% de la muestra está en esta categoría). Esperamos que las variables ficticias basadas en este indicador tengan un efecto positivo y quizás creciente sobre la probabilidad de estar parado, dado que asumimos que esos años adicionales que el individuo tardó en obtener su título se debieron a repeticiones de curso, capturando entonces la variable una falta de habilidades para el estudio que pueden ser valoradas por el mercado de trabajo. No obstante, también es posible que en estas categorías se encuentren, especialmente para la educación superior, individuos que interrumpieron su carrera educativa para trabajar, lo que no debería, en sí mismo, aumentar la propensión al desempleo.

En cuanto a las características personales, se incluyen indicadores de edad en intervalos de 4 ó 5 años, desde 16-19 (único grupo de 4 años), 20-24, ... hasta 60-64, siendo este último el grupo de referencia. Incluimos también el sexo del trabajador y un indicador para los inmigrantes, basándonos en el lugar de su nacimiento. Así, agrupamos a los inmigrantes en los siguientes grupos: UE-15 (incluyendo EEUU y Canadá), resto de Europa, Iberoamérica, África y Asia-Oceanía. Para estos individuos nacidos fuera de España, añadimos dos indicadores suplementarios de nacionalidad (española o doble), a fin de comprobar si la nacionalización conlleva algún tipo de asimilación con respecto al empleo. Finalmente, y para todos los individuos, incluimos indicadores de movilidad geográfica (cambio de residencia en el último año, bien dentro de la misma provincia, desde otra provincia o desde el extranjero).

Las características del hogar que pensamos que pueden afectar al empleo son, en primer lugar, si el individuo vive en pareja (independientemente de su estado civil), e indicadores binarios de la existencia de personas dependientes en tres categorías: niños de hasta 5 años, niños de 6 a 15 años y jubilados de 65 o más años. Las variables de niños y jubilados se obtienen contando en cada hogar a los residentes de las edades correspondientes (que además sean inactivos, en el caso de los mayores de 64).

Terminando con los regresores de nuestra especificación, intentamos capturar la variabilidad regional en el desempleo a través de distintos indicadores a nivel provincial: tasa de paro, tasa de actividad, proporción de inmigrantes y proporción de titulados universitarios entre los activos de cada provincia. Estas tasas se calculan para cada provincia (contando como tales a Ceuta y Melilla) y trimestre, siendo las únicas variables no binarias en el modelo. Finalmente, incluimos *dummies* para las 7 regiones de España definidas en el nivel NUTS-1, siendo la región 3 (Madrid) la referencia en nuestro análisis.

La Tabla 1 muestra el número de observaciones con valores válidos para todas las variables en cada periodo², así como el porcentaje de éstas que representan a parados. El crecimiento de las observaciones se debe tanto al lento crecimiento de la tasa de actividad (ver Gráfico 1) como al aumento de la muestra de la EPA en cada ciclo, asociado a la subida de la población Española.

Tabla 1: Observaciones y porcentaje de parados por trimestre

Trimestre	Obs.	% Parados	Trimestre	Obs.	% Parados
2005t1	69.481	10,38	2009t1	76.786	16,14
2005t2	68.892	9,63	2009t2	76.841	16,94
2005t3	67.182	8,71	2009t3	80.832	16,52
2005t4	67.689	9,09	2009t4	80.533	17,34
2006t1	69.771	9,31	2010t1	80.348	18,65
2006t2	72.896	8,69	2010t2	79.635	18,81
2006t3	72.287	8,27	2010t3	79.218	18,63
2006t4	72.577	8,28	2010t4	78.920	19,26
2007t1	74.677	8,46	2011t1	79.377	20,21
2007t2	76.022	8,00	2011t2	79.410	19,86
2007t3	76.074	7,98	2011t3	78.634	20,3
2007t4	76.024	8,59	2011t4	77.207	21,57
2008t1	76.702	9,43	2012t1	79.090	23,36
2008t2	77.510	10,18	2012t2	79.532	23,94
2008t3	77.271	11,02	2012t3	78.845	24,2
2008t4	76.889	13,25	2012t4	78.457	25,17

A partir de estos datos, estimamos un modelo *logit* para cada trimestre, en el que la variable dependiente $y_{i,t}$ es un indicador binario que toma el valor 1 en caso de que el individuo esté parado y 0 en caso contrario (ocupado). Para cada periodo t , la probabilidad de estar parado, condicionada a la información de cada individuo, es:

² Aunque existen algunos valores perdidos, su frecuencia es pequeña, un 0,78% en el total de la base de datos y ésta no supera el 1% en ningún trimestre.

$$\Pr(y_{i,t} = 1) = \frac{e^{x'_{i,t}\beta_t}}{1 + e^{x'_{i,t}\beta_t}}, \quad i = 1 \dots N_t, t = 1 \dots 32$$

Donde $x_{i,t}$ es un vector que recoge la información del individuo i de la muestra t y β_t es un vector de parámetros que puede variar para cada periodo. Alternativamente, puede concebirse que el modelo consiste de una variable continua latente, $y_{i,t}^* = x'_{i,t}\beta_t + u_t$ que, en caso de tomar valores positivos, hace que la variable observada $y_{i,t}$ valga 1. En tal caso, al adoptar la especificación logística estamos asumiendo que las perturbaciones u_t siguen una distribución logística estándar. Escogemos esta especificación frente a la distribución normal (lo que llevaría a un modelo *probit*) por la mayor facilidad de interpretar los coeficientes en β_t , como el logaritmo del cambio en la probabilidad relativa (*odds ratio*) al cambiar la variable explicativa³.

La estrategia empírica consiste, como se ha dicho, en estimar un modelo con los datos de corte transversal de cada ciclo de la EPA. Además, para facilitar la interpretación de los resultados, calculamos los efectos marginales de las distintas variables. Dado que la mayor parte del modelo se basa en variables cualitativas, siendo los regresores grupos de variables dicotómicas, estos efectos marginales se calculan como el cambio en la probabilidad de estar parado al pasar la variable correspondiente de la categoría de referencia a la categoría para la que se calcula el efecto marginal, manteniendo el resto de variables en sus niveles medios.

3 Resultados

A fin de sintetizar, presentamos en esta sección los resultados completos del modelo sólo para dos periodos: los segundos trimestres de 2007, por ser el periodo con menor tasa de paro, y de 2012, por ser el último trimestre comparable en nuestra serie⁴, lo que hacemos en el primer sub-epígrafe. Para el resto de periodos, presentaremos, en la segunda sub-sección, gráficos con la evolución de los efectos marginales de algunas de las variables más relevantes.

³ McDonald y Worswick (1997) estiman, con datos canadienses, una especificación probit similar a la aquí presentada, centrándose en la probabilidad de desempleo de trabajadores inmigrantes.

⁴ Los resultados detallados de las estimaciones y los efectos marginales quedan a disposición del lector que los solicite.

3.1 Modelos en el pico de empleo y situación actual

Las Tablas 2 y 3 recogen los resultados de sendos modelos *logit* para la probabilidad individual de encontrarse en paro, frente a la alternativa de estar ocupado, en el segundo trimestre de 2007 y en el periodo equivalente de 2012. Es decir, comparamos el funcionamiento del mercado de trabajo en su zenit de actividad frente al mínimo de empleo más reciente. Aún así, puede decirse que en una primera inspección los dos modelos presentan resultados parecidos, siendo la mayoría de las variables significativas en ambos modelos con los mismos signos. A partir del valor del pseudo- R^2 y la función de verosimilitud en la que se basa, sin embargo, parece que el modelo para 2012 presenta un mejor ajuste, lo que implicaría que la crisis ha ido provocando una mayor segmentación en el mercado de trabajo, en la medida en que las mismas características demográficas condicionan con más fuerza la situación laboral en la actualidad que en los periodos de bonanza.

Tabla 2: Resultados del modelo logit para la probabilidad de paro. 2º trimestre 2007.					
	Coef.	t-stat	Ef. Marginal		Err. Std.
Educación (ref. sin primaria)					
Estudios primarios completos	-0,5486	-7,84	-0,0593	***	0,0085
Secundaria de primera etapa	-0,7335	-10,63	-0,0749	***	0,0084
Garantía social	-0,8449	-1,94	-0,0833	***	0,0322
Bachillerato	-0,9839	-13,08	-0,0928	***	0,0086
FP Grado medio	-0,9381	-11,75	-0,0898	***	0,0088
FP Grado superior	-1,1670	-13,90	-0,1039	***	0,0088
Diplomaturas	-1,4245	-16,74	-0,1169	***	0,0086
Licenciaturas y grados	-1,4973	-17,46	-0,1201	***	0,0086
Cursos de especialización superior	-1,4877	-6,07	-0,1197	***	0,0131
Doctorado / Máster univ.	-2,0563	-6,28	-0,1390	***	0,0117
Repetidor					
1	-0,0039	-0,10	-0,0003		0,0025
2	0,0782	1,46	0,0054		0,0038
3	0,2983	5,39	0,0225	***	0,0045
Grupo edad (ref. 60 a 65 años)					
16 a 20	1,8449	17,77	0,1771	***	0,0099
20 a 24	1,0848	11,22	0,0769	***	0,0058
25 a 29	0,6588	6,87	0,0389	***	0,0050
30 a 34	0,5103	5,31	0,0283	***	0,0048
35 a 39	0,4608	4,79	0,0250	***	0,0047
40 a 44	0,3536	3,70	0,0183	***	0,0046
45 a 49	0,2489	2,61	0,0123	***	0,0045
50 a 54	0,2598	2,69	0,0129	***	0,0046
55 a 59	0,2662	2,66	0,0133	***	0,0048
Pareja	-0,3318	-9,14	-0,0232	***	0,0026
Mujer	0,7931	27,81	0,0558	***	0,0020
Inmigrante (ref: nativo)					
UE-15	0,3956	3,06	0,0305	***	0,0114
Resto Europa	0,4694	5,07	0,0372	***	0,0085

Tabla 2: Resultados del modelo logit para la probabilidad de paro. 2º trimestre 2007.					
Iberoamérica	0,3771	5,59	0,0288	***	0,0058
África	0,7561	7,80	0,0665	***	0,0107
Asia y Oceanía	0,2286	0,84	0,0165		0,0214
Nacionalidad si inmigrante (ref: extranjera)					
Española	-0,1573	-1,33	-0,0102		0,0072
Doble nacionalidad	-0,0069	-0,05	-0,0005		0,0102
Niño/s <5 en hogar	-0,0292	-0,66	-0,0020		0,0030
Niño/s 6 a 15 años en hogar	0,0659	1,96	0,0046	*	0,0024
Jubilado/s en hogar	0,0790	1,92	0,0055	*	0,0030
Región (NUTS-1)					
Noroeste	0,1202	1,05	0,0081		0,0075
Noreste	0,0406	0,41	0,0026		0,0064
Centro	0,0841	0,74	0,0056		0,0074
Este	0,0848	0,81	0,0056		0,0068
Sur	0,0734	0,63	0,0048		0,0075
Canarias	0,2923	2,48	0,0210	***	0,0081
Mov. geográfica último año					
Desde otro municipio de la provincia	0,5377	4,38	0,0447	***	0,0122
Desde otra provincial	0,5470	3,99	0,0456	***	0,0137
Desde el extranjero	0,6928	4,87	0,0609	***	0,0156
Variables provinciales					
Tasa de paro	11,0496	15,21	0,7593	***	0,0501
Tasa de actividad	-0,4315	-0,81	-0,0297		0,0367
Proporción de inmigrantes en la pob. activa	-0,8286	-2,53	-0,0569	**	0,0225
Proporción de universitarios en la pob. act.	1,2125	2,23	0,0833	**	0,0374
Constante	-3,3545	-9,61			
N. obs: 76.022; Log. Verosimilitud -19252,108 Pseudo R^2 : 0,0913; Test Wald (χ^2 , 47 g.l.) 4000,10 ***					

*** Coeficiente/Efecto marginal significativo al 1%; ** al 5%; * al 10%.

En términos de la magnitud de los efectos marginales, se observa que los mayores corresponden a las variables de edad y educativas, observándose en ambos casos un perfil decreciente de la probabilidad de encontrarse parado a medida que aumentan estas variables. Así, los coeficientes negativos y significativos en el caso de las variables educativas, de mayor magnitud para los niveles de educación superior, reflejan una menor probabilidad de sufrir paro en comparación con el grupo de referencia (trabajadores sin estudios primarios completos). Para las variables de edad, la referencia son los individuos de entre 60 y 64 años, siendo menos probable para el resto encontrar un empleo.

Para el caso de la educación, los resultados pueden explicarse con la teoría del capital humano (Becker, 1964), según la cual una mayor educación implica una mayor productividad. Ante un mercado de trabajo con exceso de oferta, parece lógico que las empresas prefieran contratar a individuos más productivos. Sin embargo, también debe tenerse en cuenta que la educación tiende a aumentar los salarios (Becker, 1964; Caparrós y otros, 2010 muestran evidencia reciente para España), por lo que el efecto de la educación es doble, aumentando la

probabilidad de estar empleado y el salario en caso de estarlo. En cuanto a nuestro indicador sobre repetición, cabe señalar que éste es significativo, al menos en su tercer grado (en general, tener una edad unos 4-5 años por encima de la normal al terminar cada nivel de estudios), lo que puede indicar que los estudiantes especialmente malos tienen mayores dificultades para encontrar empleo, por la señal adversa que emiten a los empleadores.

Tabla 3: Resultados del modelo logit para la probabilidad de paro. 2º trimestre 2012					
	Coef.	t-stat	Ef. Marginal		Err. Std.
Educación (ref. sin primaria)					
Estudios primarios completos	-0,4214	-7,27	-0,0918	***	0,0129
Secundaria de primera etapa	-0,7913	-13,92	-0,1654	***	0,0126
Garantía social	-0,6040	-2,08	-0,1292	**	0,0580
Bachillerato	-1,3142	-21,7	-0,2534	***	0,0129
FP Grado medio	-1,1072	-17,9	-0,2210	***	0,0132
FP Grado superior	-1,4091	-22,26	-0,2672	***	0,0131
Diplomaturas	-1,7317	-26,62	-0,3088	***	0,0130
Licenciaturas y grados	-1,9509	-29,97	-0,3328	***	0,0128
Cursos de especialización superior	-2,7343	-12,04	-0,3937	***	0,0175
Doctorado / Máster univ.	-1,8004	-17,77	-0,3167	***	0,0156
Repetidor					
1	0,0233	0,94	0,0036		0,0039
2	0,2425	6,84	0,0397	***	0,0060
3	0,2681	7,37	0,0441	***	0,0062
Grupo edad (ref. 60 a 65 años)					
16 a 20	2,2822	26,78	0,4212	***	0,0164
20 a 24	1,5771	26,65	0,2646	***	0,0089
25 a 29	1,0943	19,37	0,1658	***	0,0076
30 a 34	0,8889	15,93	0,1279	***	0,0071
35 a 39	0,6962	12,47	0,0951	***	0,0069
40 a 44	0,6149	11,11	0,0821	***	0,0067
45 a 49	0,4923	8,99	0,0634	***	0,0065
50 a 54	0,4491	8,19	0,0571	***	0,0065
55 a 59	0,3348	5,9	0,0411	***	0,0067
Pareja	-0,4769	-20,8	-0,0788	***	0,0039
Mujer	0,1600	8,83	0,0254	***	0,0029
Inmigrante (ref. nativo)					
UE-15	0,2649	3,15	0,0435	***	0,0145
Resto Europa	0,7794	13,74	0,1406	***	0,0114
Iberoamérica	0,5233	11,51	0,0904	***	0,0085
África	0,9626	15,79	0,1783	***	0,0127
Asia y Oceanía	-0,8981	-4,75	-0,1111	***	0,0178
Nacionalidad si inmigrante (ref: extranjera)					
Española	-0,1726	-2,46	-0,0263	**	0,0103
Doble nacionalidad	0,0759	0,99	0,0122		0,0126
Niño/s <5 en hogar	0,0230	0,84	0,0037		0,0044
Niño/s 6 a 15 años en hogar	0,1243	5,63	0,0199	***	0,0036
Jubilado/s en hogar	0,0888	3,27	0,0143	***	0,0044
Región (NUTS-1)					
Noroeste	0,0728	0,98	0,0115		0,0116
Noreste	0,0064	0,1	0,0010		0,0102
Centro	0,0421	0,57	0,0066		0,0114
Este	0,0624	0,96	0,0098		0,0101
Sur	0,0213	0,29	0,0033		0,0116
Canarias	0,1571	1,99	0,0252	**	0,0125
Mov. geográfica último año					

Tabla 3: Resultados del modelo logit para la probabilidad de paro. 2º trimestre 2012					
Desde otro municipio de la provincia	0,0233	0,19	0,0037		0,0191
Desde otra provincial	0,5866	4,46	0,1038	***	0,0255
Desde el extranjero	0,7082	2,96	0,1277	***	0,0477
Variables provinciales					
Tasa de paro	5,1190	20,04	0,8106	***	0,0402
Tasa de actividad	-0,1145	-0,32	-0,0181		0,0561
Proporción de inmigrantes en la pob. activa	-0,6296	-2,98	-0,0997	***	0,0335
Proporción de universitarios en la pob. act.	1,3790	4,29	0,2184	***	0,0509
Constante	-2,2437	-8,75			
N. obs: 79.523; Log. Verosimilitud -38814,52; Pseudo R^2 : 0,1132; Test Wald (χ^2 , 47 g.l.) 8508,40 ***					

*** Coeficiente/Efecto marginal significativo al 1%; ** al 5%; * al 10%.

Por lo que respecta a la influencia de la edad en la probabilidad de paro es posible verla también como un reflejo del capital humano específico, en la medida en que los individuos de mayor edad han pasado, presumiblemente, más tiempo en el mercado de trabajo, acumulando habilidades laborales. Cabe matizar que la edad incluye tanto tiempo pasado en el mercado de trabajo como periodos intermedios de desempleo y actividad, de modo que este efecto puede deberse no sólo a una acumulación de aptitudes, sino también a la segmentación del mercado de trabajo español que protege más a los trabajadores más estables y con más antigüedad en la empresa, que suelen tener más edad (ver, por ejemplo, Conde y otros, 2011; Bentolila y otros, 2012). Finalmente, el efecto puede deberse sencillamente a que los trabajadores de más edad han tenido más tiempo para encontrar un empleo en el que están bien emparejados y son productivos, por lo que la probabilidad de que la relación laboral se rompa es más pequeña.

En cuanto a otros determinantes del desempleo, los resultados son los esperados. En ambos casos, las mujeres tienen una menor probabilidad de encontrar empleo, con el resto de características constantes, bien por diferencias sistemáticas no observadas en el capital humano acumulado (por ejemplo, es común que, para la misma edad, la experiencia laboral de una mujer sea menor) o bien por la existencia de discriminación en la contratación. Algunos tipos de responsabilidad familiar afectan a la variable dependiente: la probabilidad de paro disminuye al vivir en pareja (mayor responsabilidad), pero parece aumentar cuando hay dependientes en el hogar (jubilados o niños de 6 a 15 años), lo que puede indicar problemas de conciliación. También, en general, se observa una mayor probabilidad de paro para los nacidos en el extranjero, siendo peor la situación de los inmigrantes africanos y del resto de Europa en comparación con

iberoamericanos y aquéllos nacidos en países desarrollados (la categoría UE-15, que incluye también a Estados Unidos y Canadá). Es sorprendente el caso de los inmigrantes asiáticos⁵, que en 2007 tenían una probabilidad de paro similar a la de los nativos (coeficiente y efecto marginal no significativo) y que en 2012 presentan un efecto marginal negativo, con una probabilidad de paro inferior en 11 puntos porcentuales a la de los trabajadores españoles, lo que concuerda con la percepción de que estas minorías están especializadas con éxito en algunos sectores concretos de la economía, como el sector minorista, tanto como empleados que como empresarios.

En cuanto a las diferencias regionales en el desempleo individual, vemos que éstas pueden explicarse principalmente por diferencias en la tasa de paro provincial, en la medida en que se trata de la variable con un efecto marginal más acusado. Aún así, para un mismo nivel de paro provincial, los individuos tienen más probabilidad de estar parados en las provincias con más universitarios y menos inmigrantes. En el primer caso, podría pensarse que una fuerza de trabajo más formada compite con el resto de la población por unos empleos escasos, aunque el resultado no deja de ser contraintuitivo. Por la parte de los inmigrantes, un mayor nivel de éstos en la zona geográfica en cuestión no parece incrementar esa competencia, sino facilitar las oportunidades de empleo, quizás porque estos trabajadores son más móviles y tiendan a concentrarse en las provincias con mejores perspectivas laborales. Finalmente, una vez que se ha tenido en cuenta la heterogeneidad provincial a través de estas tasas, los efectos de las variables regionales desaparecen, salvo para el caso de Canarias, en las que es algo más probable encontrarse en situación de desempleo (unos 2 puntos porcentuales más en los dos trimestres analizados aquí).

En cuanto a la evolución de los resultados a medida que la crisis despliega sus efectos, ésta se verá en más detalle a continuación. Por el momento, sin embargo, destaca que tanto los coeficientes como los efectos marginales tienden a crecer en 2012 con respecto a 2007. Esto sucede tanto con las variables educativas como con la edad, inmigración, situación familiar y variables regionales. Esto parece indicar que, en términos de empleo, la crisis ha provocado una

⁵ El INE no proporciona datos por país, sino por regiones más amplias. De este grupo, aproximadamente la mitad proceden de Asia del sur y central, presumiblemente de India y Pakistán en su mayoría, y el resto de Asia oriental, presumiblemente de la República Popular China.

mayor segmentación en la economía española: en la actualidad, la edad, el nivel educativo y la procedencia son mucho más determinantes a la hora de conseguir empleo.

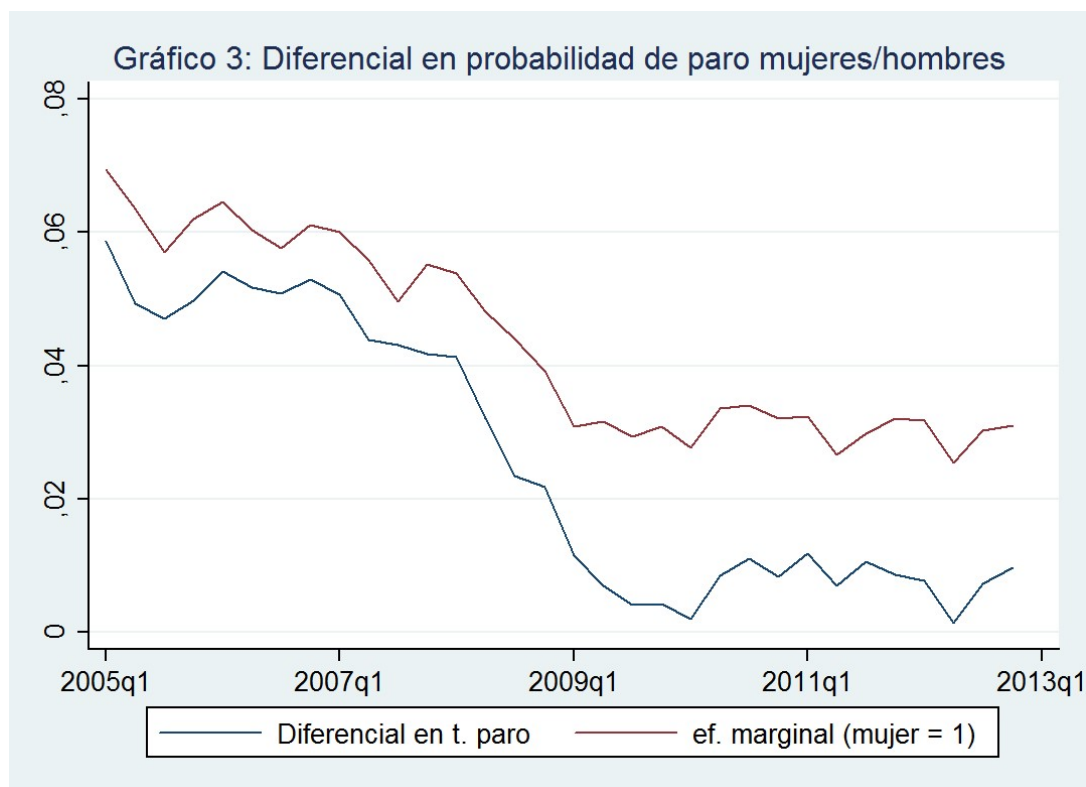
La principal excepción es el caso de las mujeres, cuyo diferencial positivo se reduce: el ratio de probabilidades relativas para las mujeres de estar en paro pasa de 2,21 a 1,17, y el efecto marginal pasa de una probabilidad un 5,6% superior de estar parada que un hombre a sólo un 2,5%, menos de la mitad. Esto es especialmente notable si tenemos en cuenta que en ambos periodos la tasa de desempleo media aumenta del 8% al 26%.

3.2 Evolución de los determinantes individuales del desempleo

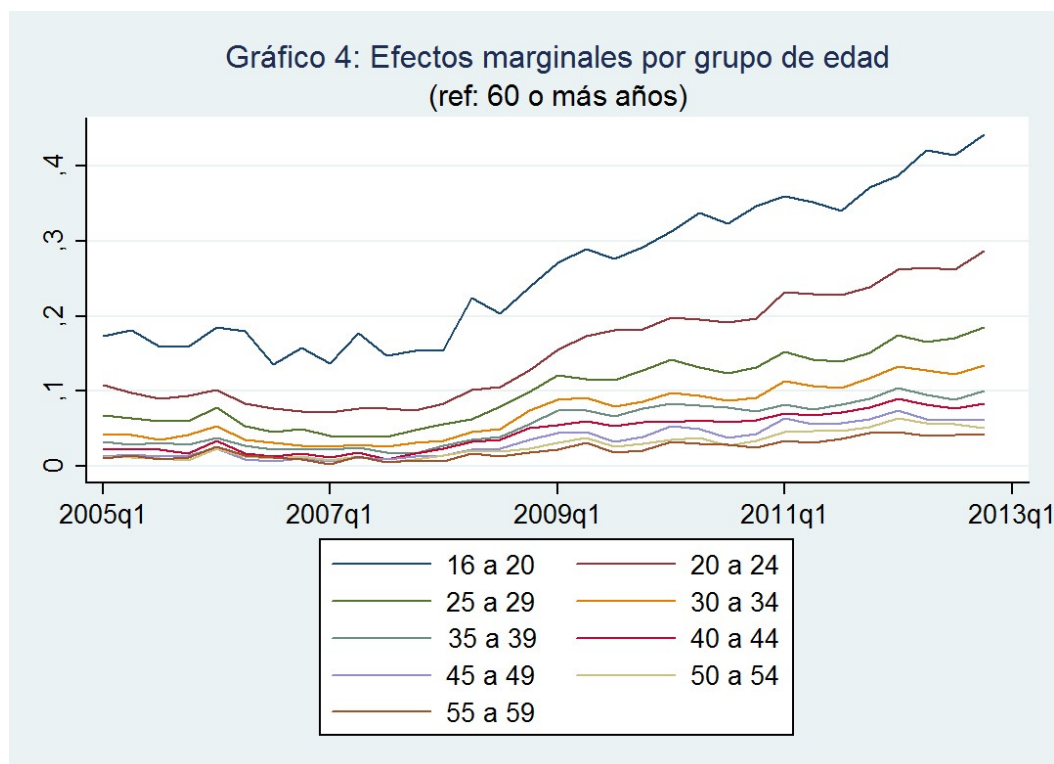
En este sub-epígrafe se presentan de forma gráfica los efectos marginales de algunas de las variables de los 32 modelos estimados (uno por cuatrimestre), viendo su evolución temporal. Así, en primer lugar, el Gráfico 3 muestra (línea roja) la evolución del efecto marginal de la variable mujer⁶, así como la diferencia observada en tasas de paro a lo largo del periodo analizado (línea azul).

Puede comprobarse que ambos indicadores disminuyen a lo largo del tiempo. Como ya se vio anteriormente, las tasas de paro “brutas” (observadas) quedan prácticamente equiparadas para ambos sexos a partir de 2009. Sin embargo, el efecto marginal, que disminuía paralelamente hasta esas fechas, se mantiene a partir de entonces en un nivel en torno al 3%. La principal explicación de esta igualación puede ser el hecho de que gran parte del empleo desaparecido durante la crisis haya sido en sectores de actividad y ocupaciones en los que los hombres estaban más especializados, especialmente la construcción.

⁶ Éste se calcula como el incremento en la probabilidad de estar parado para las mujeres, $\Pr(y_{i,t} = 1 | mujer = 1, \bar{x}_t) - \Pr(y_{i,t} = 1 | mujer = 0, \bar{x}_t)$, donde \bar{x}_t se refiere a los valores medios del resto de variables explicativas.



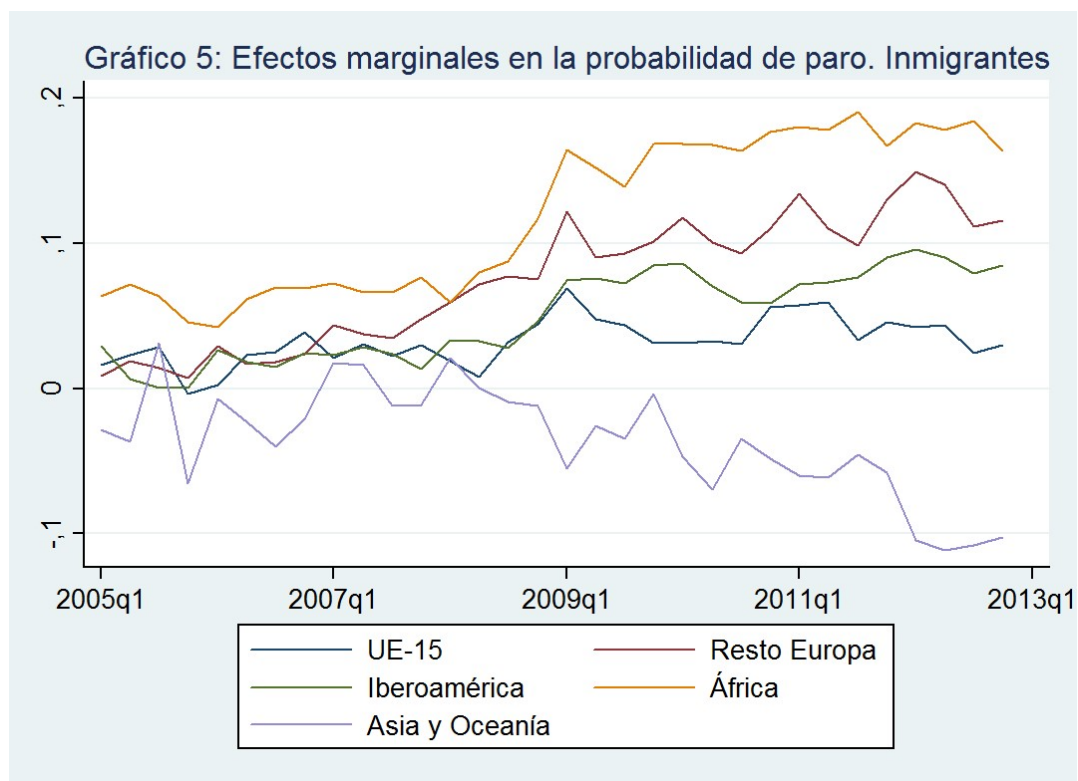
Sin embargo, ese diferencial en la probabilidad de un 3% indica que para una mujer es más difícil encontrar un empleo que para un varón equivalente. Dado que las tasas de paro están más cercanas que ese 3%, y que la diferencia siempre es menor al efecto estimado, se concluye que, en realidad, las mujeres presentan unas características personales que, en principio, deberían llevarlas a experimentar una tasa de desempleo inferior a la masculina. Pese a que la crisis ha implicado un progreso relativo (dado que, no debe olvidarse, que durante el periodo analizado la probabilidad de desempleo sube para los dos grupos), éste parece haberse estancado por alguna razón en la segunda fase de la crisis.



Siguiendo con la evolución de los determinantes del paro, pasamos a los indicadores de edad (Gráfico 4). En este caso, las líneas representan el diferencial en la probabilidad de experimentar desempleo para cada grupo de edad en comparación con el de referencia (mayores de 60 años). Puede comprobarse, que las diferencias se redujeron en la fase de bonanza, hasta 2007, donde las líneas de los trabajadores desde 40 a 59 años están cercanas a 0 (de hecho, en el modelo presentado en la Tabla 2, aunque tienen más probabilidad de estar parados que los mayores de 60 años, no hay diferencias significativas en los tramos de edad de 45 a 59 años). A partir de entonces, sin embargo, los grupos empiezan a separarse unos de otro. Cabe señalar, además, que en este gráfico los diferenciales en probabilidad son de una magnitud muy acusada, hasta de un 40% para los menores de 20 años y de entre 20 y 30 puntos para los trabajadores en la veintena (líneas verde y marrón). Esto implica un mercado de trabajo extremadamente segmentado por edad, lo que puede tener graves consecuencias para las cohortes más jóvenes, para las que es muy difícil acumular experiencia laboral y construir una carrera estable⁷.

⁷ Debe advertirse, sin embargo, que estos resultados no informan sobre las posibilidades de encontrar empleo para un parado a distintas edades, dado que la probabilidad estimada no se condiciona a un estado

En el Gráfico 5 aparece algo parecido, aunque de una dimensión relativa menor. Así, las probabilidades individuales de estar en paro estaban relativamente igualadas hasta el inicio de la crisis y, a partir de entonces, se observa una divergencia al inicio de la crisis que parece estabilizarse a partir de 2010. La peor situación es para los inmigrantes nacidos en África y países Europeos no de la UE-15 (de éstos, alrededor de dos tercios son trabajadores procedentes de Rumanía y Bulgaria, los dos últimos países en incorporarse a la Unión Europea), cuya probabilidad de estar parados supera a la de los españoles en más de un 10%.



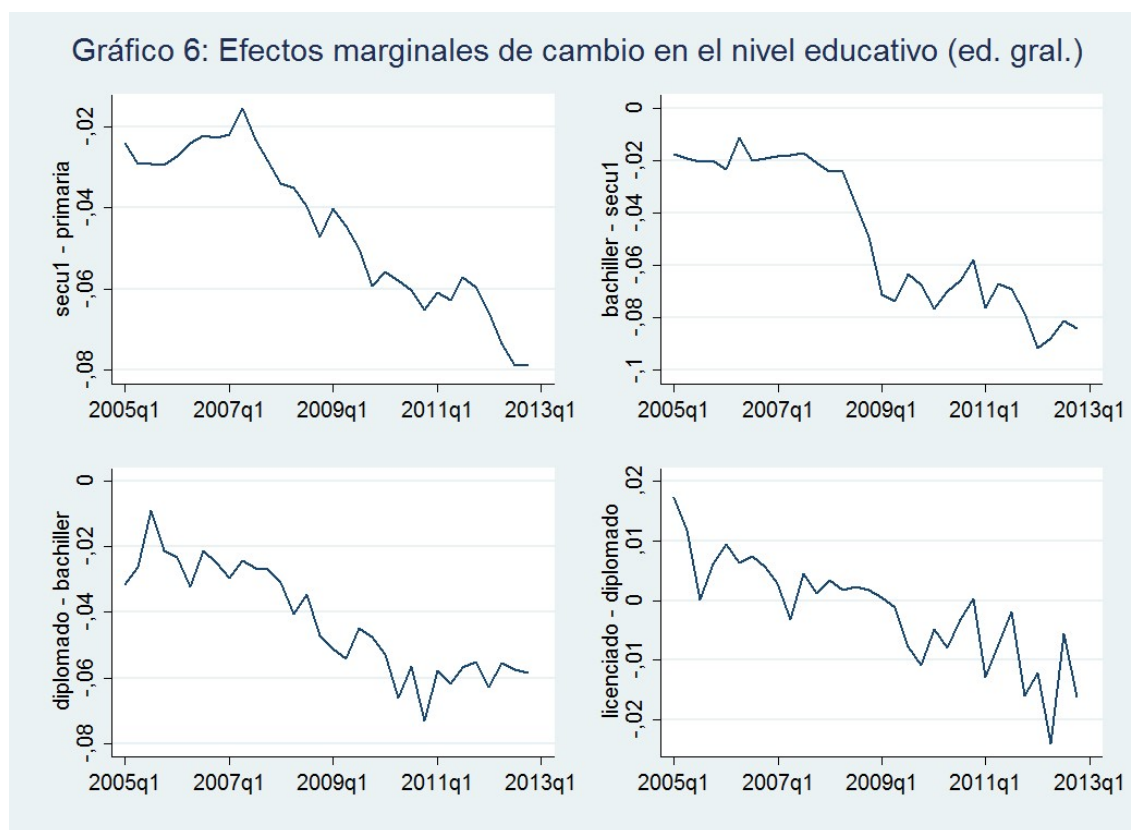
También se observa un empeoramiento, aunque menor, en las condiciones para los inmigrantes iberoamericanos y de países desarrollados (mayoritariamente UE-15). Para el caso de los trabajadores americanos, su mejor posición en relación con otros inmigrantes podría estar relacionada con el dominio del idioma, dado que la mayoría proceden de países hispano-

laboral previo. En FEDEA (2013), con datos de la EPA enlazada, se muestra que no hay grandes diferencias en la probabilidad de pasar del paro al empleo en función de la edad, salvo para los menores de 25 años, que salen perjudicados.

hablantes. En cuanto a los nacionales de la Unión Europea, parece lógico pensar que, en caso de encontrarse en situación de paro, la posibilidad de retorno a sus países de origen sería más atractiva que en otros casos. Finalmente, vuelve a resultar llamativo el buen comportamiento de los inmigrantes de Asia y Oceanía, que a partir de 2011 presentan una probabilidad significativamente menor que la de los españoles de estar en paro y que supera los 10 puntos porcentuales en los últimos trimestres disponibles. Este último resultado, sin embargo, debe relativizarse, pues estos trabajadores representan apenas el 0,4% de los activos y el 3,8% de los activos inmigrantes a finales de 2012 (aunque su presencia ha aumentado notablemente desde 2005)

Pasamos, finalmente, a analizar la evolución de la importancia de los estudios en el empleo. Así, el Gráfico 6 muestra la evolución de los efectos de ascender un “escalón” en distintos niveles de educación generalista (frente a vocacional): de primaria a secundaria de primera etapa (EGB o ESO, dependiendo de la cohorte), de secundaria 1 a bachillerato (diagrama arriba a la derecha), de bachillerato a diplomaturas y de éstas a licenciaturas (abajo a la derecha). Para interpretar los cuatro diagramas, debe señalarse que, mientras que en los tres primeros la escala del eje vertical es parecida (de 0 a -10 puntos porcentuales), ésta es mucho más reducida en el último, que compara distintos títulos universitarios (de 2 a -2 puntos porcentuales). En cualquier caso, vemos que, salvo en el caso de las licenciaturas, cada nivel educativo implica una probabilidad individual de desempleo inferior que en el nivel educativo inmediatamente anterior. Además, estas diferencias se incrementan drásticamente a partir de 2007, con el inicio de la crisis. Así, pasar desde estudios primarios al primer nivel de secundaria o de éste a bachillerato reducía la probabilidad de paro en un 2% al principio de la crisis y ahora lo hace en un 8%. En el caso de los títulos universitarios de ciclo corto (diplomaturas e ingenierías técnicas), la probabilidad de paro es ahora menor en un 6%. El caso de las licenciaturas y las diplomaturas es especial, pues al principio del periodo analizado no existían diferencias importantes (aunque sí significativas, y a favor de los diplomados, en los dos primeros trimestres de la muestra), mientras que se ha abierto

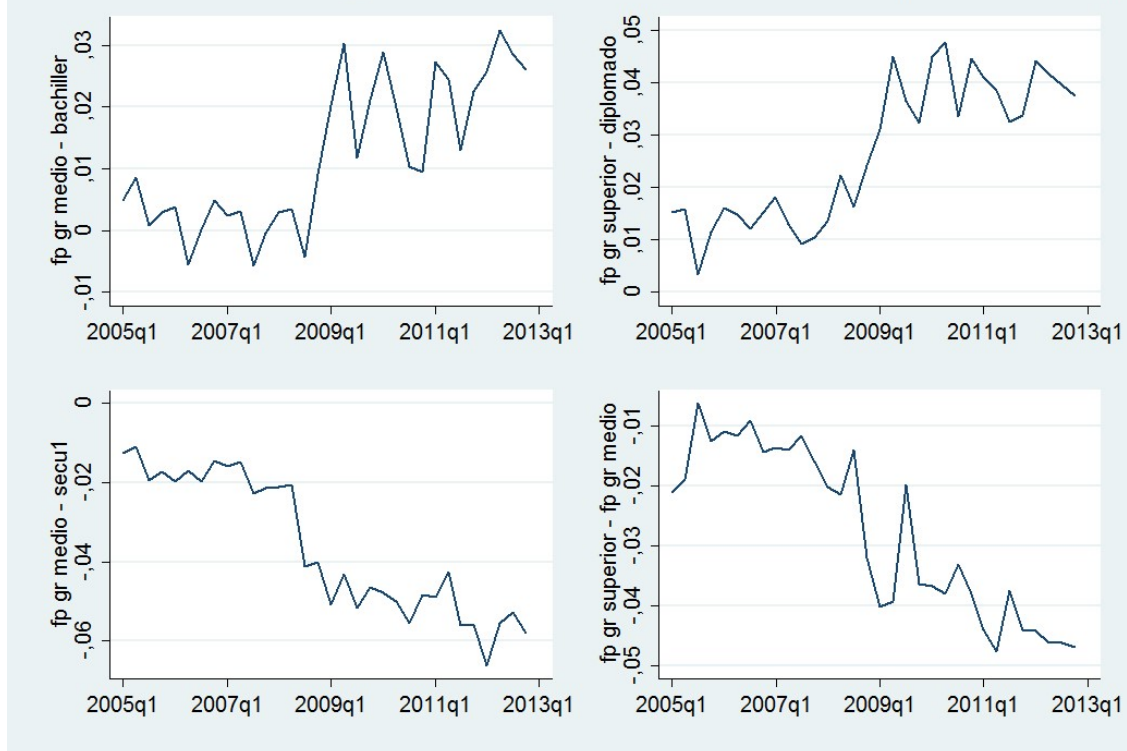
una brecha que favorece a los licenciados en los dos últimos años analizados (el efecto marginal, de en torno a 1-2 puntos porcentuales es significativo en 5 de los 8 trimestres de 2011 y 2012).



En cuanto a la educación vocacional o profesional, es interesante compararla con sus niveles “equivalentes” generalistas (esto es, bachillerato frente a formación profesional de grado medio y diplomaturas frente a formación profesional de grado superior, dado que a las dos se accede, en principio, con el título de bachillerato). Esta comparación aparece en las dos primeras figuras del Gráfico 7. En ambos casos se aprecia que la probabilidad de paro relativa aumenta con el tiempo para los titulados de formación profesional, aunque en el caso de la comparación del grado medio con bachillerato las diferencias son erráticas y no siempre significativas. Las diferencias son, en cualquier caso, moderadas, de unos 2 a 4 puntos porcentuales a partir de 2010. En principio, cabría esperar un mejor acomodo de los egresados de títulos vocacionales, de modo que estos resultados podrían estar indicando un problema en el sistema de formación profesional

español, aunque también puede deberse a que a menudo los alumnos con más dificultades se derivan a estas titulaciones.

Gráfico 7: Efectos marginales de cambio en el nivel educativo (ed. vocacional)



Por otra parte, en la parte inferior del Gráfico 7 se comparan las dos categorías de títulos profesionales frente a su alternativa inmediatamente anterior: FP de grado medio frente a secundaria obligatoria (abajo a la izquierda) y FP superior frente a grado medio (abajo a la derecha). En este caso, se parte de unas diferencias pequeñas que, de nuevo, se amplifican a partir del comienzo de la crisis. En este caso, el efecto marginal acaba siendo en 2011-2012 más grande para los titulados de grado medio, que ven reducida su probabilidad de estar en paro un 6% respecto a los titulados de secundaria. El efecto adicional para los graduados superiores parece estar estabilizado en unos 4-5 puntos porcentuales.

Finalmente, y en conjunto sobre el efecto de la educación en la probabilidad de paro, parece lógico que un mayor nivel formativo se asocie a unas mejores perspectivas de empleo. Sin embargo, la divergencia observada en prácticamente todos los niveles educativos no puede

atribuirse a una súbita mejoría del sistema educativo (puesto que la mayor parte de los individuos analizados acabaron sus estudios bien antes de comenzar la crisis), y quizás tampoco a una mayor racionalidad del mercado de trabajo, que en los años de bonanza pudo crear empleos para trabajadores de poca cualificación. Más bien, estos diferenciales son preocupantes dado que, para el nivel de paro actual, pueden llegar a excluir del empleo a los individuos menos formados e impedirles acumular experiencia, de modo que su situación se haga permanente.

4 Conclusiones

Este trabajo ha presentado evidencia sobre los determinantes individuales del desempleo en España y su evolución en el pasado reciente, con vistas a ver los efectos de la crisis en distintos grupos de trabajadores. Para ello, hemos estimado la probabilidad de estar parado (frente a ocupado) con modelos *logit* sobre los datos trimestrales de la EPA desde 2005 hasta 2012. Los resultados muestran que, a partir de la segunda mitad de 2007, se produce una clara segmentación del mercado laboral, pues las mismas características influyen en la probabilidad de estar parado pero con una intensidad mucho mayor. Esto sucede principalmente con la edad y el nivel educativo, pero también con la condición de inmigrante. Estas diferencias pueden reflejar una preferencia de los empleadores por trabajadores mejor formados y con más experiencia, lo que es razonable, pero las probabilidades de paro estimadas en los colectivos más desfavorecidos (jóvenes, inmigrantes y con nivel educativo bajo) se sitúan fácilmente por encima del 50%, lo que implica que segmentos importantes de la población pueden ser excluidos del mercado de trabajo. Aunque no se ha estudiado aquí, la inercia individual del empleo es bastante elevada (Stewart, 2007, proporciona evidencia para el Reino Unido junto con el asunto relacionado de la dependencia de estado en los empleos de bajos salarios; para España, resultados recientes aparecen en Aguilar y otros, 2012), por lo que esos individuos podrían enfrentarse a una vida laboral extremadamente intermitente e insatisfactoria.

Referencias

- Aguilar, M^a Isabel, Dolores García y M^a Luz González (2012), " Transiciones laborales en España: influencia de las características familiares". Comunicación presentada en las XXI Jornadas de la Asociación de Economía de la Educación (AEDE). Oporto, Portuga.
- Becker, Gary S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. The University of Chicago Press, Chicago. 3^a edición, 1993.
- Bentolila, Samuel, Pierre Cahuc, Juan J. Dolado y Thomas Le Barbanchon (2012), " Two-Tier labour markets in the Great Recession: France versus Spain". *The Economic Journal*, 122, pp. F155-F187.
- Caparrós, Antonio; M^a Lucía Navarro y Mario F. Rueda (2010), "Endogenous wage determinants and returns to education in Spain". *International Journal of Manpower*, 31(4), pp. 410-425.
- Conde, J. Ignacio, Florentino Felgueroso y J. Ignacio García (2011), "Reforma laboral 2010: Una primera evaluación y propuestas de mejora". *Revista de Economía Aplicada*, 57, pp. 147-180.
- Fundación de Estudios de Economía Aplicada (FEDEA, 2013). "Observatorio laboral de la crisis", Boletín electrónico XVIII.
- García-Serrano, Carlos (2011), "Déjà vu? Crisis de empleo y reformas laborales en España". *Revista de Economía Aplicada*, 56, pp. 149-177.
- McDonald, James T. y Christopher Worswick (1997), "Unemployment incidence of immigrant men in Canada". *Canadian Public Policy*, vol. 23(4), pp. 353-373.
- Stewart, Mark S. (2007), "The interrelated dynamics of unemployment and low-wage employment". *Journal of Applied Econometrics*, 22. pp. 511-531.